

ENLG 2011

# **The 13th European Workshop on Natural Language Generation**

**Proceedings**

28 – 30 September 2011

ENLG 2011 is sponsored by



©2011 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)  
209 N. Eighth Street  
Stroudsburg, PA 18360  
USA  
Tel: +1-570-476-8006  
Fax: +1-570-476-0860  
[acl@aclweb.org](mailto:acl@aclweb.org)

## Introduction

It is with great pleasure that we present the current volume of papers accepted for presentation at the 13th European Workshop on Natural Language Generation (ENLG 2011), which will be held from September 28th to 30th, 2011 at Loria in Nancy, France.

The ENLG 2009 workshop continued a biennial series of workshops on natural language generation that has been running since 1987 and alternates with INLG, the International Conference on Natural Language Generation. Previous European workshops have been held at Royaumont, Edinburgh, Judenstein, Pisa, Leiden, Duisburg, Toulouse, Budapest, Aberdeen, Dagstuhl and Athens. Together with INLG, the ENLG workshop is the main regular forum for presenting and discussing research in Natural Language Generation.

ENLG 2011 invited submissions on all topics related to natural language generation. We received 41 submissions of long and short papers from all over the world. Of these 13 long papers and 12 short papers were accepted for presentation. The long papers will be presented orally, and the short papers as posters.

In addition, ENLG 2011 hosts Generation Challenges 2011. This year, three shared task evaluation competitions were organized under the umbrella of Generation Challenges 2011: the Surface Realisation Challenge (Belz, Hogan, White, and Stent), the Challenge on Generating Instructions in Virtual Environments (Striegnitz, Denis, Gargett, Garoufi, Koller, and Theune) and the Helping Our Own Challenge (Dale and Kilgariff).

The first part of this volume contains the 25 research papers that will be presented at ENLG 2011. The second part is devoted to the Generation Challenges 2011 session. It contains overview reports on the active and planned challenges and system descriptions of all participating teams.

We are indebted to the authors and to the members of our program committee whose hard work contributed to making this a collection of high quality research papers. We are also delighted that Oliver Lemon, Johanna Moore and Jeff Orkin agreed to give invited talks at ENLG 2011. And last but not least, many thanks go to the local organisation team, Nicolas Alcaraz, Anne Lise Charbonnier, Alexandre Denis, Alejandra Lorenzo, Shashi Narayan and Laura Perez-Beltrachini for handling the preparation of the meeting.

Claire Gardent and Kristina Striegnitz  
Program co-Chairs for ENLG 2011



**Program Co-Chairs:**

Claire Gardent, CNRS/Loria, Nancy, France  
Kristina Striegnitz, Union College, USA

**Local Organizers:**

Nicolas Alcaraz, INRIA Nancy, France  
Anne-Lise Charbonnier, INRIA Nancy, France  
Alexandre Denis, CNRS/LORIA, Nancy, France  
Alejandra Lorenzo, INRIA/LORIA, Nancy, France  
Shashi Narayan, U. Nancy/LORIA, Nancy, France  
Laura Perez-Beltrachini, U. Nancy/LORIA, Nancy, France

**Program Committee:**

John Bateman, University Bremen, Germany  
Anja Belz, University of Brighton, UK  
Bernd Bohnet, University Stuttgart, Germany  
Stephan Busemann, DFKI, Germany  
Christian Chiarcos, University of Potsdam, Germany  
Norman Creaney, University of Ulster, Ireland  
Robert Dale, Macquarie University, Australia  
Kees van Deemter, University of Aberdeen, Scotland  
Seniz Demir, Tubitak-Bilgem, Turkey  
Alexandre Denis, CNRS/LORIA Nancy, France  
David DeVault, USC Institute for Creative Technologies, USA  
Barbara Di Eugenio, University of Illinois, USA  
Roger Evans, University of Brighton, UK  
Leo Ferres, University of Concepcion, Chile  
Jennifer Foster, Dublin University, Ireland  
Albert Gatt, University of Malta, Malta  
Josef van Genabith, Dublin City University, Ireland  
Pablo Gervas, Universidad Complutense de Madrid, Spain  
Markus Guhe, University of Edinburgh, UK  
John Kelleher, Dublin Institute of Technology, Ireland  
Alistair Knott, University of Otago, New Zealand  
Alexander Koller, University of Saarbrcken, Germany  
Stefan Kopp, University of Bielefeld, Germany  
Eric Kow, University of Brighton, UK  
Emiel Krahmer, Tilburg University, The Netherlands  
Geert-Jan Kruijff, DFKI, Germany  
Ivana Kruijff-Korabayova, DFKI, Germany  
Oliver Lemon, Heriot Watt University, Edinburgh, Scotland

James Lester, North Carolina State University, USA  
Keith van der Linden, Calvin College, USA  
François Mairesse, University of Cambridge, UK  
Kathleen McCoy, University of Delaware, USA  
David McDonald, SIFT, Inc., USA  
Chris Mellish, University of Aberdeen, Scotland  
Jon Oberlander, University of Edinburgh, Scotland  
Cécile Paris, CSIRO ICT Centre, Australia  
Paul Piwek, The Open University, UK  
Richard Power, The Open University, UK  
Ehud Reiter, University of Aberdeen and Data2Text Ltd, Scotland  
Donia Scott, University of Sussex, UK  
Advaith Siddharthan, University of Aberdeen, Scotland  
Ielka van der Sluis, Trinity College Dublin, Ireland  
Yaji Sripada, University of Aberdeen, Scotland  
Manfred Stede, University of Potsdam, Germany  
Amanda Stent, AT&T Labs Research, USA  
Matthew Stone, Rutgers, USA  
Michael Strube, EML Research, Germany  
Mariët Theune, University of Twente, The Netherlands  
Takenobu Tokugana, Tokyo Institute of Technology, Japan  
Jette Viethen, Macquarie University, Australia  
Carl Vogel, Trinity College Dublin, Ireland  
Michael White, Ohio State University, USA  
Sandra Williams, the Open University, UK  
Tie-Jun Zhao, Harbin Institute of Technology, China  
Michael Zock, CNRS/LIF Université de la Méditerranée Aix-Marseille II, France

**Invited Speakers:**

Oliver Lemon, Heriot Watt University, Scotland  
Johanna Moore, University of Edinburgh, Scotland  
Jeff Orkin, MIT Media Lab, USA

# Table of Contents

## Oral Presentations

<i>Talkin' bout a revolution (statistically speaking)</i> (Invited Talk)	
Oliver Lemon .....	1
<i>Text Simplification using Typed Dependencies: A Comparision of the Robustness of Different Generation Strategies</i>	
Advait Siddharthan .....	2
<i>Generating Affective Natural Language for Parents of Neonatal Infants</i>	
Saad Mahamood and Ehud Reiter .....	12
<i>What is in a text and what does it do: Qualitative Evaluations of an NLG system – the BT-Nurse – using content analysis and discourse analysis</i>	
Rahul Sambaraju, Ehud Reiter, Robert Logie, Andy Mckinlay, Chris McVittie, Albert Gatt and Cindy Sykes .....	22
<i>Evaluating Saliency Metrics for the Context-Adequate Realization of Discourse Referents</i>	
Christian Chiarcos .....	32
<i>The Impact of Visual Context on the Content of Referring Expressions</i>	
Henriette Viethen, Robert Dale and Markus Guhe .....	44
<i>A Cross-Linguistic Study on the Production of Multimodal Referring Expressions in Dialogue</i>	
Ielka Van Der Sluis and Saturnino Luz .....	53
<i>Two Approaches for Generating Size Modifiers</i>	
Margaret Mitchell, Kees Van Deemter and Ehud Reiter .....	63
<i>Using Online Games to Capture, Generate, and Understand Natural Language</i> (Invited Talk)	
Jeff Orkin .....	71
<i>Content selection from an ontology-based knowledge base for the generation of football summaries</i>	
Nadjet Bouayad-Agha, Gerard Casamayor and Leo Wanner .....	72
<i>Deriving rhetorical relationships from semantic content</i>	
Richard Power .....	82
<i>If it may have happened before, it happened, but not necessarily before</i>	
Albert Gatt and François Portet .....	91
<i>Adaptive Information Presentation for Spoken Dialogue Systems: Evaluation with real users</i>	
Verena Rieser, Simon Keizer, Oliver Lemon and Xingkun Liu .....	102

<i>Combining Hierarchical Reinforcement Learning and Bayesian Networks for Natural Language Generation in Situated Dialogue</i>	
Nina Dethlefs and Heriberto Cuayáhuatl .....	110
<i>Combining symbolic and corpus-based approaches for the generation of successful referring expressions</i>	
Konstantina Garoufi and Alexander Koller .....	121
<i>Language Generation for Spoken Dialogue Systems</i> (Invited Talk)	
Johanna D. Moore .....	132
<b>Poster Presentations</b>	
<i>Adapting SimpleNLG to German</i>	
Marcel Bollmann .....	133
<i>EasyText: an Operational NLG System</i>	
Laurence Danlos, Frédéric Meunier and Vanessa Combet .....	139
<i>Towards Generating Text from Discourse Representation Structures</i>	
Valerio Basile and Johan Bos .....	145
<i>A Policy-Based Approach to Context Dependent Natural Language Generation</i>	
Thomas Bouttat, Edoardo Pignotti, Chris Mellish and Peter Edwards .....	151
<i>Levels of organisation in ontology verbalisation</i>	
Sandra Williams, Allan Third and Richard Power .....	158
<i>Using semantic roles to improve summaries</i>	
Diana Trandabâţ .....	164
<i>Building a Generator for Italian Sign Language</i>	
Alessandro Mazzei .....	170
<i>Investigation into Human Preference between Common and Unambiguous Lexical Substitutions</i>	
Andrew Walker, Advaith Siddharthan and Andrew Starkey .....	176
<i>Production of Demonstratives in Dutch, English and Portuguese Dialogues</i>	
Saturnino Luz and Ielka Van Der Sluis .....	181
<i>Generation of Formal and Informal Sentences</i>	
Fadi Abu Sheikha and Diana Inkpen .....	187
<i>Glue Rules for Robust Chart Realization</i>	
Michael White .....	194
<i>Detecting Interesting Event Sequences for Sports Reporting</i>	
François Lareau, Mark Dras and Robert Dale .....	200



## Generation Challenges 2011

<i>Generation Challenges 2011 Preface</i>	
Anja Belz, Albert Gatt, Alexander Koller and Kristina Striegnitz .....	206
<i>The GRUVE Challenge: Generating Routes under Uncertainty in Virtual Environments</i>	
Srini Janarthanam and Oliver Lemon .....	208
<i>A Proposal for a Spanish Surface Realization Shared Task</i>	
Pablo Gervás and Miguel Ballesteros .....	212
<i>The First Surface Realisation Shared Task: Overview and Evaluation Results</i>	
Anja Belz, Mike White, Dominic Espinosa, Eric Kow, Deirdre Hogan and Amanda Stent . . .	217
<i>DCU at Generation Challenges 2011 Surface Realisation Track</i>	
Yuqing Guo, Deirdre Hogan and Josef van Genabith .....	227
<i>ATT-0: Submission to Generation Challenges 2011 Surface Realization Shared Task</i>	
Amanda Stent .....	230
<i>&lt;StuMaBa &gt;: From Deep Representation to Surface</i>	
Bernd Bohnet, Simon Mille, Benoît Favre and Leo Wanner .....	232
<i>The OSU System for Surface Realization at Generation Challenges 2011</i>	
Rajakrishnan Rajkumar, Dominic Espinosa and Michael White .....	236
<i>UCM Submission to the Surface Realization Challenge</i>	
Pablo Gervás .....	239
<i>Helping Our Own: The HOO 2011 Pilot Shared Task</i>	
Robert Dale and Adam Kilgariff .....	242
<i>May I check the English of your paper!!!</i>	
Pinaki Bhaskar, Aniruddha Ghosh, Santanu Pal and Sivaji Bandyopadhyay .....	250
<i>Handling Outlandish Occurrences: Using Rules and Lexicons for Correcting NLP Articles</i>	
Elitza Ivanova, Delphine Bernhard and Cyril Grouin .....	254
<i>NUS at the HOO 2011 Pilot Shared Task</i>	
Daniel Dahlmeier, Hwee Tou Ng and Thanh Phu Tran .....	257
<i>Helping Our Own 2011: UKP Lab System Description</i>	
Torsten Zesch .....	260
<i>University of Illinois System in HOO Text Correction Shared Task</i>	
Alla Rozovskaya, Mark Sammons, Joshua Gioja and Dan Roth .....	263
<i>Data-Driven Correction of FunctionWords in Non-Native English</i>	
Adriane Boyd and Detmar Meurers .....	267

<i>Report on the Second Second Challenge on Generating Instructions in Virtual Environments (GIVE-2.5)</i>	
Kristina Striegnitz, Alexandre Denis, Andrew Gargett, Konstantina Garoufi, Alexander Koller and Mariët Theune . . . . .	270
<i>Direction giving: an attempt to increase user engagement</i>	
Bob Duncan and Kees van Deemter . . . . .	280
<i>The Bremen System for the GIVE-2.5 Challenge</i>	
Nina Dethlefs . . . . .	284
<i>The GIVE-2.5 C Generation System</i>	
David Nicolás Racca, Luciana Benotti and Pablo Duboue . . . . .	290
<i>CL system: Giving instructions by corpus based selection</i>	
Luciana Benotti and Alexandre Denis . . . . .	296
<i>The Loria Instruction Generation System L in GIVE 2.5</i>	
Alexandre Denis . . . . .	302
<i>The Potsdam NLG systems at the GIVE-2.5 Challenge</i>	
Konstantina Garoufi and Alexander Koller . . . . .	307
<i>The Thumbs Up! Twente system for GIVE 2.5</i>	
Saskia Akkersdijk, Marin Langenbach, Frieder Loch and Mariët Theune . . . . .	312
<i>Question Generation Shared Task and Evaluation Challenge – Status Report</i>	
Vasile Rus, Brendan Wyse, Paul Piwek, Mihai Lintean, Svetlana Stoyanchev and Cristian Moldovan . . . . .	318

# Conference Program

## Wednesday, September 28, 2011

8:30–9:15      Registration

9:15–9:30      Opening Remarks

9:30–10:30     Invited Talk

*Talkin' bout a revolution (statistically speaking)*

Oliver Lemon

10:30–11:00    Coffee Break

### **ENLG Talks: Tailoring text to users**

11:00–11:30    *Text Simplification using Typed Dependencies: A Comparison of the Robustness of Different Generation Strategies*

Advaith Siddharthan

11:30–12:00    *Generating Affective Natural Language for Parents of Neonatal Infants*

Saad Mahamood and Ehud Reiter

12:00–12:30    *What is in a text and what does it do: Qualitative Evaluations of an NLG system – the BT-Nurse – using content analysis and discourse analysis*

Rahul Sambaraju, Ehud Reiter, Robert Logie, Andy Mckinlay, Chris McVittie, Albert Gatt and Cindy Sykes

12:30–1:30     Lunch

### **ENLG Talks: Referring expression generation**

1:30–2:00      *Evaluating Saliency Metrics for the Context-Adequate Realization of Discourse Referrers*

Christian Chiarcos

2:00–2:30      *The Impact of Visual Context on the Content of Referring Expressions*

Henriette Viethen, Robert Dale and Markus Guhe

2:30–3:00      Coffee Break

**Wednesday, September 28, 2011 (continued)**

**ENLG Talks: Referring expression generation**

- 3:00–3:30     *A Cross-Linguistic Study on the Production of Multimodal Referring Expressions in Dialogue*  
Ielka Van Der Sluis and Saturnino Luz
- 3:30–4:00     *Two Approaches for Generating Size Modifiers*  
Margaret Mitchell, Kees Van Deemter and Ehud Reiter
- 1:30-2:30     Birds-of-a-Feather sessions / SimpleNLG User Group Meeting
- 6:30-8:00     Bowling and Drinks

**Thursday, September 29, 2011**

- 9:30–10:30     Invited Talk

*Using Online Games to Capture, Generate, and Understand Natural Language*  
Jeff Orkin

- 10:30–11:00     Coffee Break

**ENLG Talks: Knowledge representation and NLG — expressing semantic, rhetorical and temporal relations**

- 11:00–11:30     *Content selection from an ontology-based knowledge base for the generation of football summaries*  
Nadjet Bouayad-Agha, Gerard Casamayor and Leo Wanner
- 11:30–12:00     *Deriving rhetorical relationships from semantic content*  
Richard Power
- 12:00–12:30     *If it may have happened before, it happened, but not necessarily before*  
Albert Gatt and François Portet
- 12:30–1:30     Lunch

## Thursday, September 29, 2011 (continued)

### Generation Challenges Talks and Posters

- |           |   |
|-----------|---|
| 1:30–1:40 | Generation Challenges Introductory Remarks  |
| 1:40–2:55 | Result Presentation: Surface Realization (SR), Helping Our Own (HOO), Generating Instructions in Virtual Environments (GIVE)                                |
| 2:55–3:40 | Presentation of Planned Tasks: SR Spanish, Question Generation 2nd Edition, Generating Route Instructions under Uncertainty in Virtual Environments (GRUVE) |
| 3:40–5:00 | Generation Challenges Poster Session (with coffee)  |
| 6:00–7:00 | Guided Tour in the Old Town   |
| 7:30      | Banquet   |

## Friday, September 30, 2011

### ENLG Talks: Optimizing task success in interactive systems

- |             |   |
|-------------|---|
| 9:30–10:00  | <i>Adaptive Information Presentation for Spoken Dialogue Systems: Evaluation with real users</i><br>Verena Rieser, Simon Keizer, Oliver Lemon and Xingkun Liu             |
| 10:00–10:30 | <i>Combining Hierarchical Reinforcement Learning and Bayesian Networks for Natural Language Generation in Situated Dialogue</i><br>Nina Dethlefs and Heriberto Cuayáhuatl |
| 10:30–11:00 | <i>Combining symbolic and corpus-based approaches for the generation of successful referring expressions</i><br>Konstantina Garoufi and Alexander Koller                  |
| 11:00–12:30 | ENLG Posters (see below for a detailed list)  |
| 12:30–1:30  | Lunch   |

**Friday, September 30, 2011 (continued)**

- 1:30–2:30      Invited Talk
- Language Generation for Spoken Dialogue Systems*  
Johanna D. Moore
- 2:30–3:00      Coffee Break
- 3:00–5:00      Generation Challenges Working Groups
- 5:00–5:30      Reports from the Working Groups and Closing Remarks

**ENLG Poster Presentations**

*Adapting SimpleNLG to German*  
Marcel Bollmann

*EasyText: an Operational NLG System*  
Laurence Danlos, Frédéric Meunier and Vanessa Combet

*Towards Generating Text from Discourse Representation Structures*  
Valerio Basile and Johan Bos

*A Policy-Based Approach to Context Dependent Natural Language Generation*  
Thomas Bouttaz, Edoardo Pignotti, Chris Mellish and Peter Edwards

*Levels of organisation in ontology verbalisation*  
Sandra Williams, Allan Third and Richard Power

*Using semantic roles to improve summaries*  
Diana Trandabăt

*Building a Generator for Italian Sign Language*  
Alessandro Mazzei

*Investigation into Human Preference between Common and Unambiguous Lexical Substitutions*  
Andrew Walker, Advait Siddharthan and Andrew Starkey

**Friday, September 30, 2011 (continued)**

*Production of Demonstratives in Dutch, English and Portuguese Dialogues*

Saturnino Luz and Ielka Van Der Sluis

*Generation of Formal and Informal Sentences*

Fadi Abu Sheikha and Diana Inkpen

*Glue Rules for Robust Chart Realization*

Michael White

*Detecting Interesting Event Sequences for Sports Reporting*

François Lareau, Mark Dras and Robert Dale





# Using Online Games to Capture, Generate, and Understand Natural Language

**Jeff Orkin**

MIT Media Laboratory

Cambridge, MA, USA

jorkin@media.mit.edu

While the game industry has excelled at simulating combat, dynamically generating social interaction and natural language dialogue has proven intractable. The gaming landscape today, with millions of people playing games together online, provides opportunities to radically rethink our approach to developing conversational, socially intelligent characters. This talk will present a data-driven, human-machine collaborative approach to automating characters using data recorded from online human-human interactions, including a crowd-sourced annotation framework, and a new real-time planning system driven by thousands of annotated human gameplay traces. The approach will be demonstrated with examples from three novel games: The Restaurant Game has recorded over 16,000 people playing as customers and waitress, Improviso is currently recording players on the set of a low-budget science fiction movie, and Mars Escape has recorded hundreds of online human-robot interactions for eventual transfer to a real robot.